

SEQUENCE LISTING

<110> BENOIT, Patrick
SCHWARTZ, Bertrand
BRANELLEC, Didier
CHIEN, Kenneth R.

<120> SEQUENCES UPSTREAM OF THE CARP GENE, VECTORS CONTAINING
THEM AND USES THEREOF

<130> 03806.0530-00000

<140>

<141> 2001-12-07

<160> 5

<170> PatentIn Ver. 2.1

<210> 1

<211> 2358

<212> DNA

<213> Mus musculus

<400> 1

```

ggatcctttc atgtttaaca atatcaacc taacccaagg ggaacagcct gcctgacagt 60
ggctttgcca cccatgaata cttcctagtc tagtccgttt gtgaaactca gcccatccca 120
acacttctgc aagccccatc ctctacaagg tgctcattgg gaatttcctg gagcttctct 180
ttcaggatca gcctgattct agggcagcag ttctcaacct gggggcctcg acccctttgg 240
gggaatcaaa cgacccttta caggggtcac atatcatcta tcctatatgt caggtattta 300
cattacgatt cgtaacagta gcaaaattac aggtatgaaa tagcaatgaa ataattttat 360
gattgaaggc caccacaaca tgaggccgcc acactgttct agagaaaaat cacctgggtg 420
gggaaagggt tgggaaagcc tttctgtcca ttcttcattc ttcaaagtga tgtgttcaca 480
gaaagccttt cagctgttct gctggggctc ttagtaagtc tgagtaggaa ctgtatgtac 540
caggtctgct tcttatgggt ggagccaaga cgcacgtgg gtggagcgaa gacgcaacct 600
caccttctag ctctgcatcc atagcaagta gcctaattgt tctgtgtcta ggtgtcatct 660
ctgtgaatcg agatccttgg ccttgcttga attagggagg cacaaaatac tcagagattc 720

```

aagactgctc agcagcccag agtccttcct caaaggaaag gtctcaactc tcagcccccc 780
ttagctctga gtcaggcctg gaacaaacgg ccacaggaat gagaaaagct gccatagctg 840
cttgtcactt caagaggtca aagaaaatag tgттаaccat gaaaacgaga agaccaacag 900
ttatccattg atagcgtctc aggacagata ggacagagag aacactagga gaggggaacc 960
cacgaaggac aaggtattag tgtgttggtt ttcagggcaa tgtcttgtag tgaagattct 1020
agaaacacaa tttgctgggt gaacagctga agtgggggtg gggttcttac cccatgttca 1080
tggaaggggtg agtgaggaga gacagatata tgatggccag cataacaaac atacacaaca 1140
ccctaattaa cacttcctc ttctactgac acccccttca ctctcctctt tcataaaaaa 1200
taaaaaaagt attttatgtg gctcttacga tagaatcttt cctcgaacta taaaaagatc 1260
taaatattta tatttttcac attttaatat cttagcgatg acaagccaga aacaagtatt 1320
ttttgcctct ctcaacagca aagcttgggg cttttttggt tccgtgttag gaatagaaca 1380
cgagagcccc gtgtatctag gcagatgctc tatcattagc ccatgagtct ccagcctcag 1440
acgcacattt ttctcgggct ctcttaagct tttcccacag cattgggaaa ctttactgac 1500
agcatccaag ttgtgcttct gctaagaact ggactcacat ctctctgtgc atcacttcgg 1560
cccgttttgg ggtagatcct ctgattagcc ttcagattta gaacacgggtg agcctgtggt 1620
gcactaatta tggccagtga caccatagag tcaaagtga ttactgaatg ctttcaattt 1680
ctcctaattg tggtacgatg gcatgtcaca gggccatttt agctgcagac atcactccag 1740
agaattccaa acagatagag acaagtggca cccagacca tctccttccc ctcgggctga 1800
ttatccccag aaataggatg tcccaaagca acacttccca gccaaactgga gtgctgataa 1860
gtccagttat cagaaagata tggctgtaag tgtgatgcac agtgcttgca ttttcttgat 1920
acgttagtca tatgagagct gacaaagaag gaaaaagagc agcgatgtgg tgcaatatta 1980
acaggcagct gtcccctggc ttcccagatac gtgggatgac tcgcattgct gagcgggtgtg 2040
gtcactgcca aaggaatgac cctctcacat ttcttcctga ttcgcatacg ccgcggccag 2100
cttgtcatct ccctcttggg cttcccagac actaagtctg gaatgaaaat tcacctgcct 2160
ctgaattggc cactgggtggg ggcaggggtg tgacttggtt tcccaggctg gaagattatc 2220
tcaccagcc ctagctatat aacgggctgg tgtggagggg ctccacaggg ccagttccag 2280
gggttcatcc acaagagaga aaaacataga ctcgaggtct agggagcttg catgcctgca 2340

ggtcggaggc caccatgg

2358

<210> 2
<211> 2074
<212> DNA
<213> Homo sapiens

<400> 2

ctgcagcaag ttacttaatg ttttttgcct cagcatcctc tctgtaaaat gagagcatta 60
gtcttgctcc aacttcgagg gcatggacag ctctgggatt tcatatccaa gacccttaaa 120
catcccacag tccttcccc aaacacttct cctcctaata cctccctcag tttgggtcag 180
gcctggaaca aaaaggcata cgaaatggta gaaaaagtgt ccatgactac ttctgactta 240
gatgaagaga ccaatgaaaa tagtaatgac tctgtttgct tcagcaggac atatactaaa 300
ataggagcta tacaagaag attagcatgg actctgtgca agaatgacac acaaatttgt 360
gaaacattcc atatattaaa aataaataaa taataaagag aaaaggaaaa aattaaag 420
aaaatagtga tagctgtgtc catctcaaag aaaagcccag gagatttcct ttatttacct 480
cctttaagat agaatattag gagaccggaa catatgatac aggaggtact gggagggtcc 540
ctctttgtca atgttttgtc ttgggggtggg gagtcgatgt cttctcaaag tttcagaaac 600
accatccact gactgagcat tcaaggggca agaggagaat ggcagccaca tttgttgatt 660
gggtgagttt ggggagaaat agacacacaa aggtcaaaca taacttccta attaacactt 720
ccctccattc acaattccct tctcccattc ttctctcctg tcttttacts akaraaacc 780
agtttttcct gaaactataa aaatacccc agtatgttta cataatttac acctcaaaga 840
ttagaaacca gaaatagaga ctttttcaac cttccggaa gcaaagtgc ttatccctcc 900
agccacgtgt ctcaaattct gatgcatcag aatcatctgg gtgctttkaa attcaagatg 960
attcctacga gttaccataa atcaactcag aattccctgg agtggggcca gggatctgta 1020
tttctgacaa gctcccacag gtgattcctt tccccacagc atttgagaac ttcagctcaa 1080
tgacctaatc agagtcctgc cattgcta atctgggtctc atttttbtca tatatatata 1140
tagtatttgt ggtagagatg ggattttgcc atgttgccca ggctagtatt gaactcctaa 1200
gctaagcaat cttcctgtct ctgcctcca aaatgttggg attacaggtg taagccactg 1260

caccggctg atagctggtt tcatttactc tatttcttga ccactctgat ccattttgaa 1320
 gtaaaaatgc tccaattatt atgctgtttt agaacacggt aagcatgtca tgtgctaattg 1380
 gccagtgaca tcataaaaga aaagtgcatt actgaatgct ttcaatgtct tataatgatg 1440
 gtaaggtggc atgtcatggg gcctatttag ccagacatc actccaaaga attccaaaca 1500
 gatatagaca agtgccttta gggcccagat cccttcccct caggctgttt acccagggaa 1560
 taggatgtcc tgggacaagt tttccctaag tgaagtgttg ataagtctgc ttatcagaaa 1620
 gatattactg ggggtgtgat atgtagggca tctacatttt cttgataggt agtcatatga 1680
 aagctgacaa agaaaaaaag ggcagtgatg tgggtgcaatg tcaacagaca gctgtcccct 1740
 gactcttgac aaataggatg acttgcattg ctgagcgatg tgatcaccac caaaggaatg 1800
 gccctctcac atttcttcct gattcacata ttcagcaggg ttagcttgtc ctcccctccc 1860
 tcttcagctt ccagacact gagtctggaa tgaaaattca cctgcctctg agttggctcc 1920
 taatgggggc gggagtgtta cttcggttcc cagggttgaa gattatctca cccggcccca 1980
 gctatataag ctgaccggtg tggagggggc cagcagggcc aactccaggg attccttcca 2040
 cgacagaaaa acatacaaga ctccttcagc caac 2074

<210> 3
 <211> 750
 <212> PRT
 <213> Homo sapiens

<400> 3
 Met Gly Glu Thr Leu Gly Asp Ser Pro Ile Asp Pro Glu Ser Asp Ser
 1 5 10 15
 Phe Thr Asp Thr Leu Ser Ala Asn Ile Ser Gln Glu Met Thr Met Val
 20 25 30
 Asp Thr Glu Met Pro Phe Trp Pro Thr Asn Phe Gly Ile Ser Ser Val
 35 40 45
 Asp Leu Ser Val Met Glu Asp His Ser His Ser Phe Asp Ile Lys Pro
 50 55 60
 Phe Thr Thr Val Asp Phe Ser Ser Ile Ser Thr Pro His Tyr Glu Asp
 65 70 75 80
 Ile Pro Phe Thr Arg Thr Asp Pro Val Val Ala Asp Tyr Lys Tyr Asp
 85 90 95

Leu Lys Leu Gln Glu Tyr Gln Ser Ala Ile Lys Val Glu Pro Ala Ser
100 105 110

Pro Pro Tyr Tyr Ser Glu Lys Thr Gln Leu Tyr Asn Lys Pro His Glu
115 120 125

Glu Pro Ser Asn Ser Leu Met Ala Ile Glu Cys Arg Val Cys Gly Asp
130 135 140

Lys Ala Ser Gly Phe His Tyr Gly Val His Ala Cys Glu Gly Cys Lys
145 150 155 160

Gly Phe Phe Arg Arg Thr Ile Arg Leu Lys Leu Ile Tyr Asp Arg Cys
165 170 175

Asp Leu Asn Cys Arg Ile His Lys Lys Ser Arg Asn Lys Cys Gln Tyr
180 185 190

Cys Arg Phe Gln Lys Cys Leu Ala Val Gly Met Ser His Asn Ala Ile
195 200 205

Arg Phe Gly Arg Met Pro Gln Ala Glu Lys Glu Lys Leu Leu Ala Glu
210 215 220

Ile Ser Ser Asp Ile Asp Gln Leu Asn Pro Glu Ser Ala Asp Leu Arg
225 230 235 240

Ala Leu Ala Lys His Leu Tyr Asp Ser Tyr Ile Lys Ser Phe Pro Leu
245 250 255

Thr Lys Ala Lys Ala Arg Ala Ile Leu Thr Gly Lys Thr Thr Asp Lys
260 265 270

Ser Pro Phe Val Ile Tyr Asp Met Asn Ser Leu Met Met Gly Glu Asp
275 280 285

Lys Ile Lys Phe Lys His Ile Thr Pro Leu Gln Glu Gln Ser Lys Glu
290 295 300

Val Ala Ile Arg Ile Phe Gln Gly Cys Gln Phe Arg Ser Val Glu Ala
305 310 315 320

Val Gln Glu Ile Thr Glu Tyr Ala Lys Ser Ile Pro Gly Phe Val Asn
325 330 335

Leu Asp Leu Asn Asp Gln Val Thr Leu Leu Lys Tyr Gly Val His Glu
340 345 350

Ile Ile Tyr Thr Met Leu Ala Ser Leu Met Asn Lys Asp Gly Val Leu
355 360 365

Ile Ser Glu Gly Gln Gly Phe Met Thr Arg Glu Phe Leu Lys Ser Leu
5

370					375					380					
Arg 385	Lys	Pro	Phe	Gly	Asp 390	Phe	Met	Glu	Pro	Lys 395	Phe	Glu	Phe	Ala	Val 400
Lys	Phe	Asn	Ala	Leu 405	Glu	Leu	Asp	Asp	Ser 410	Asp	Leu	Ala	Ile	Phe 415	Ile
Ala	Val	Ile	Ile 420	Leu	Ser	Gly	Asp	Arg 425	Pro	Gly	Leu	Leu	Asn 430	Val	Lys
Pro	Ile	Glu 435	Asp	Ile	Gln	Asp	Asn 440	Leu	Leu	Gln	Ala	Leu 445	Glu	Leu	Gln
Leu	Lys 450	Leu	Asn	His	Pro	Glu 455	Ser	Ser	Gln	Leu	Phe 460	Ala	Lys	Leu	Leu
Gln 465	Lys	Met	Thr	Asp	Leu 470	Arg	Gln	Ile	Val	Thr 475	Glu	His	Val	Gln	Leu 480
Leu	Gln	Val	Ile	Lys 485	Lys	Thr	Glu	Thr	Asp 490	Met	Ser	Leu	His	Pro 495	Leu
Leu	Gln	Glu	Ile 500	Tyr	Lys	Asp	Leu	Tyr 505	Ala	Trp	Ala	Ile	Leu 510	Thr	Gly
Lys	Thr	Thr 515	Asp	Lys	Ser	Pro	Phe 520	Val	Ile	Tyr	Asp	Met 525	Asn	Ser	Leu
Met	Met 530	Gly	Glu	Asp	Lys	Ile 535	Lys	Phe	Lys	His	Ile 540	Thr	Pro	Leu	Gln
Glu 545	Gln	Ser	Lys	Glu	Val 550	Ala	Ile	Arg	Ile	Phe 555	Gln	Gly	Cys	Gln	Phe 560
Arg	Ser	Val	Glu	Ala 565	Val	Gln	Glu	Ile	Thr 570	Glu	Tyr	Ala	Lys	Ser 575	Ile
Pro	Gly	Phe 580	Val	Asn	Leu	Asp	Leu	Asn 585	Asp	Gln	Val	Thr	Leu 590	Leu	Lys
Tyr	Gly 595	Val	His	Glu	Ile	Ile	Tyr 600	Thr	Met	Leu	Ala	Ser 605	Leu	Met	Asn
Lys	Asp 610	Gly	Val	Leu	Ile	Ser 615	Glu	Gly	Gln	Gly	Phe 620	Met	Thr	Arg	Glu
Phe 625	Leu	Lys	Ser	Leu	Arg 630	Lys	Pro	Phe	Gly	Asp 635	Phe	Met	Glu	Pro	Lys 640
Phe	Glu	Phe	Ala	Val 645	Lys	Phe	Asn	Ala	Leu 650	Glu	Leu	Asp	Asp	Ser 655	Asp

Leu Ala Ile Phe Ile Ala Val Ile Ile Leu Ser Gly Asp Arg Pro Gly
 660 665 670
 Leu Leu Asn Val Lys Pro Ile Glu Asp Ile Gln Asp Asn Leu Leu Gln
 675 680 685
 Ala Leu Glu Leu Gln Leu Lys Leu Asn His Pro Glu Ser Ser Gln Leu
 690 695 700
 Phe Ala Lys Leu Leu Gln Lys Met Thr Asp Leu Arg Gln Ile Val Thr
 705 710 715 720
 Glu His Val Gln Leu Leu Gln Val Ile Lys Lys Thr Glu Thr Asp Met
 725 730 735
 Ser Leu His Pro Leu Leu Gln Glu Ile Tyr Lys Asp Leu Tyr
 740 745 750

<210> 4
 <211> 30
 <212> DNA
 <213> Artificial sequence
 <400> 4
 GGCGATTAA ATAATGTAGT CTTATGCAAT

30

<210> 5
 <211> 31
 <212> DNA
 <213> Artificial sequence
 <400> 5
 GGGGTCTAGA AGGTGCACAC CAATGTGGTG A

31